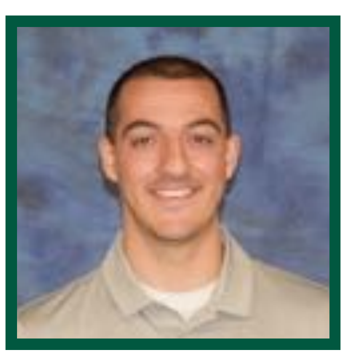
Logical shadow tomography: Efficient estimation of error-mitigated observables

Hong-Ye Hu

In collaboration with:



Ryan LaRose

Yi-Zhuang You

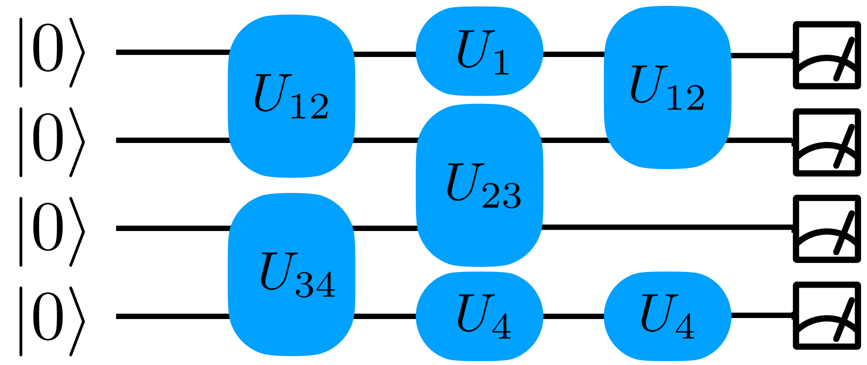
Eleanor Rieffel

Zhihui Wang

Quantum error mitigation

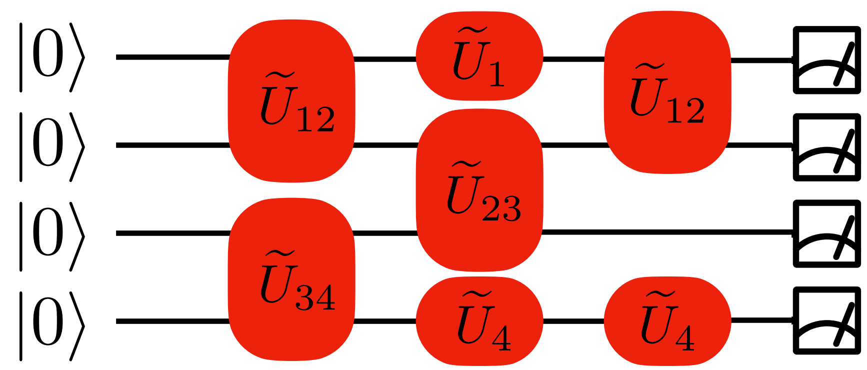
- Ideal quantum computation prepares state $\rho = |\psi\rangle\langle\psi|$, and we care about

$$\langle O \rangle_{\text{ideal}} = \text{Tr}(\rho O)$$



- Real devices have noise/decoherence, so the expectation value changes

$$\langle O \rangle_{\text{noisy}} = \text{Tr}\left(\rho_{\mathcal{E}} O\right)$$



- Error mitigation aims to get closer prediction w.r.t ideal state:

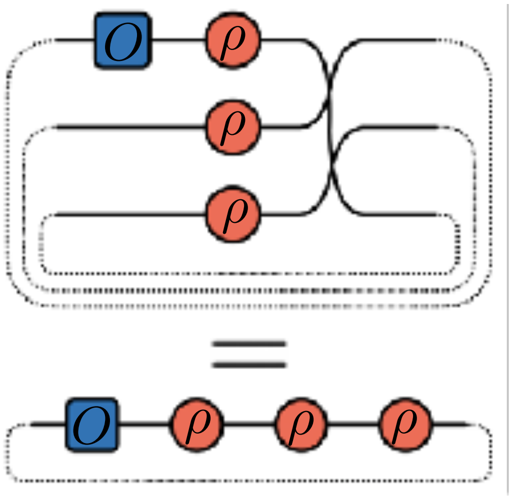
$$|\langle O \rangle_{\text{QEM}} - \langle O \rangle_{\text{ideal}}| < |\langle O \rangle_{\text{noisy}} - \langle O \rangle_{\text{ideal}}|$$

and it can be performed **on real device** or **in post-processing**.

Quantum error mitigation

- Some existing error mitigation method:
  1. Zero noise extrapolation (assume constant noise)
  2. Probabilistic error cancellation (request process tomography)

- Can we use **intermediate scale resources** fruitfully without employing the machinery of fault tolerance?



  3. Virtual distillation (VD): (request multiple copies of states)

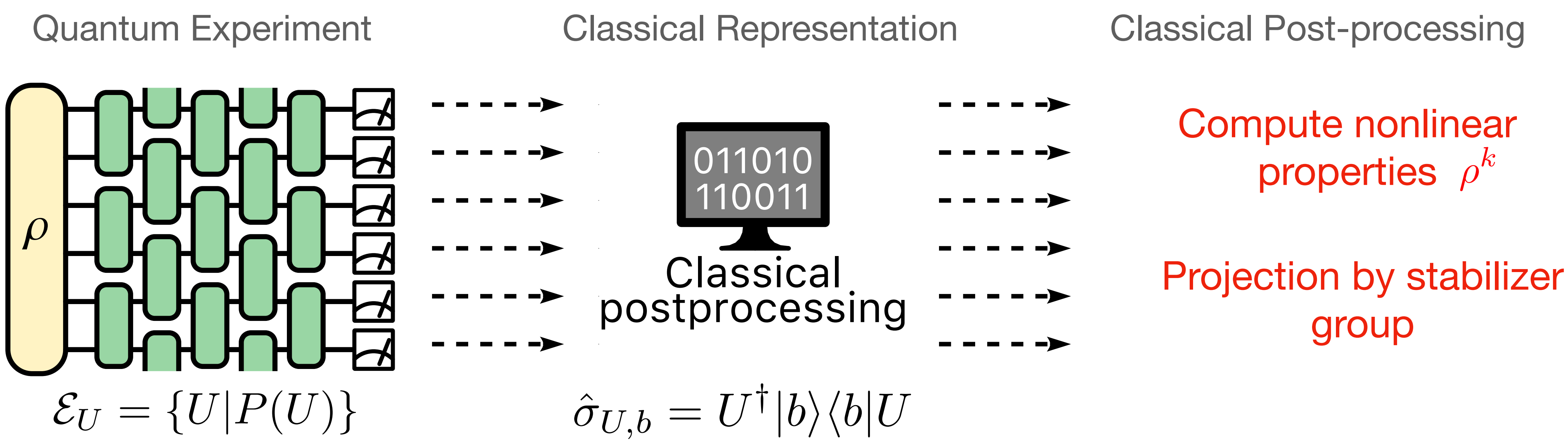
$$\langle O \rangle_{\mathrm{QEM}} = \frac{\mathrm{Tr}(O\rho^M)}{\mathrm{Tr}(\rho^M)}$$

  4. Subspace expansion (SE): error correction code w/o active correction (exponential many measurements)

     - Projection to logical subspace: $P = \Pi_{i=1}^{m} \frac{\mathbb{I}+S_i}{2} = \frac{1}{2^m} \sum_{M_i \in \mathcal{S}} M_i$
     - Projecting out non-logical error: $\langle O \rangle = \frac{1}{c2^m} \sum_i \mathrm{Tr}\left(\rho_{\mathcal{E}} M_i O\right) \quad c = \mathrm{Tr}\left(P \rho_{\mathcal{E}} P\right)$

Classical shadow representation

- Key observation: both SE and VD can be efficiently evaluated with classical shadow tomography in the post-processing phase
- Classical shadow representation:



  - Dataset: $\{\hat{\sigma}_{U,b} = U^\dagger|b\rangle\langle b|U\}$ with $P(\hat{\sigma}_{U,b}|\rho) = P(U)\mathrm{Tr}(\hat{\sigma}_{U,b}\rho)$
  - Reconstructing density matrix: $\rho = \underset{\hat{\sigma}\in\mathcal{E}_{\sigma|\rho}}{\mathbb{E}} \mathcal{M}^{-1}[\hat{\sigma}]$
- Any quantum state can be represented by a linear combination of stabilizer states

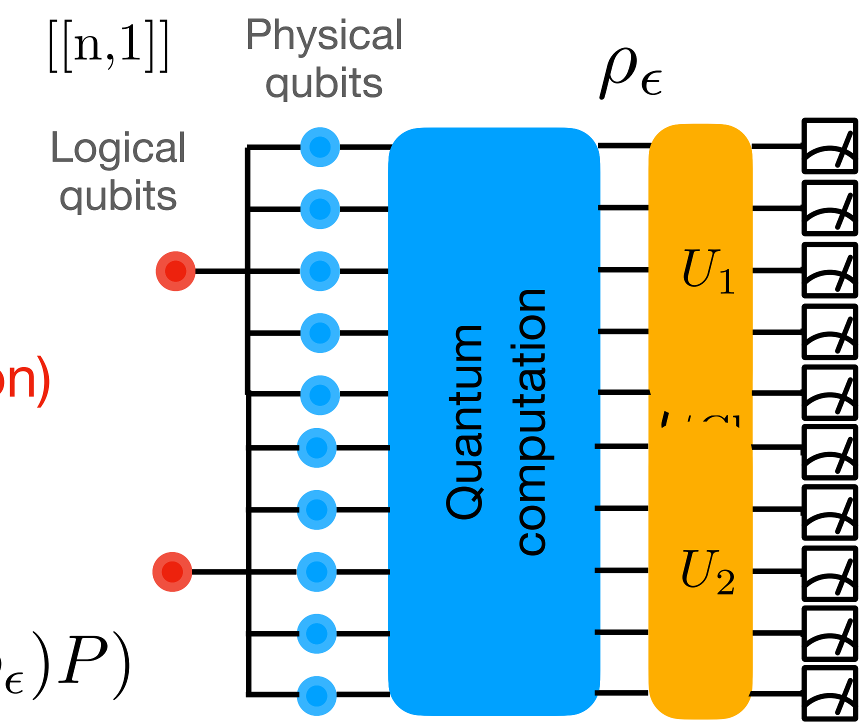
Logical shadow tomography

- General framework:
  - Error correction code defines logical subspace: $P = \Pi_{i=1}^{m} \frac{\mathbb{I} + S_i}{2}$ (no active error correction)
  - Logical shadow tomography:

    $\langle O \rangle_{\mathrm{LST}} = \frac{1}{c}\mathrm{Tr}(Pf(\rho_\epsilon)PO)$ where $c = \mathrm{Tr}(Pf(\rho_\epsilon)P)$

    In general, $f(\rho_\epsilon) = c_0 I + c_1\rho_\epsilon + c_2\rho_\epsilon^2 + \ldots$

    SE & VD can be viewed as special cases.



  - Classical shadow representation: the error mitigated value can be efficiently calculated in post-processing

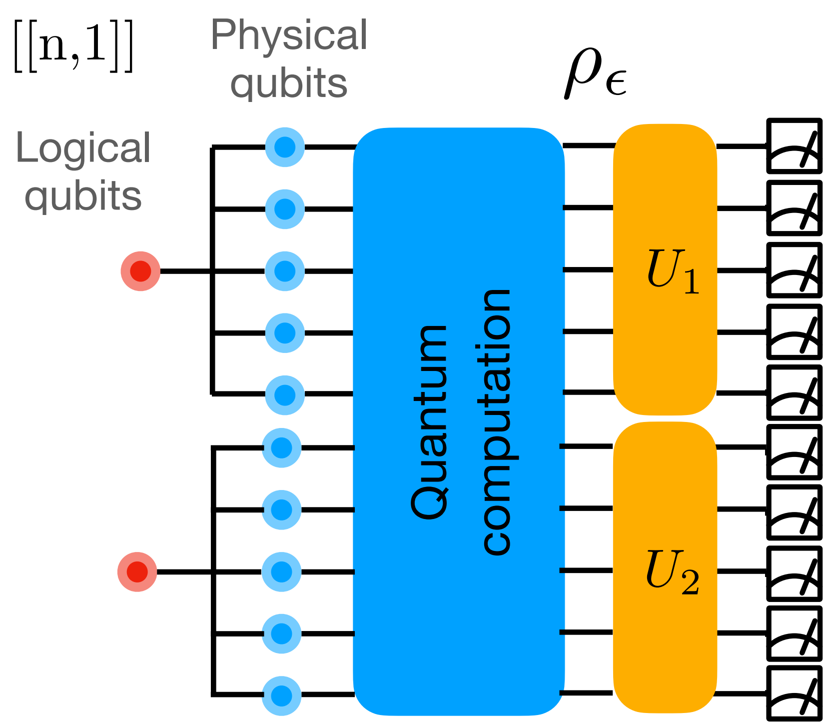
$$\langle O \rangle_{\mathrm{LST}} \approx \frac{1}{c}\mathrm{Tr}(Pf'(\hat{\sigma})PO)$$

  - What is the quantum resources and classical resources needed?

Quantum resources

- Gate overhead:

  1. Requires **Clifford** encoding circuit (same as SE)
  2. Does not require multiple copies of system (better than VD)
  3. Requires addition O(n) depth **Clifford** circuits, which doesn't scale with logical qubit number



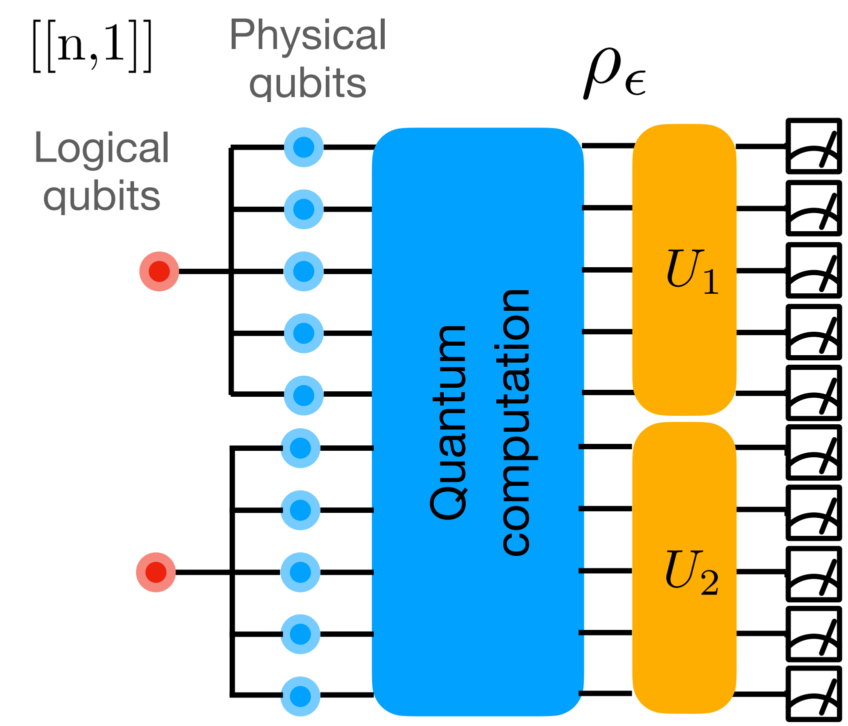
- Sample complexity:

  1. Requires $O(4^k)$ measurements to estimate logical Pauli observable $\mathrm{Tr}(P\rho PO)$

  (Better than SE: requires $O(2^{(n-1)k})$ measurements for the same task.)

- Logical shadow tomography is **better** or the **same** as VD or SE in terms of quantum resources.

Classical resources



- The estimation of $\mathrm{Tr}(P\rho^m PO)$ with classical shadows:
  1. Since each logical qubit is encoded with [[n,1]] code, the projection can be factorized:
  $$P = P_1 \otimes P_2 \otimes \cdots \otimes P_k$$
  2. The reconstruction map is also factorized:
  $$\hat{\rho} = \otimes_{i=1}^k \left((2^n+1)\hat{\sigma}_i - \mathbb{1}\right)$$
  3. Nonlinear property estimation:
  $$\mathrm{Tr}\left(P\rho^m PO\right) \approx \prod_{i=1}^k \sum_{q=0}^m \binom{m}{q} (2^n+1)^q(-1)^{m-q}\mathbb{E}_{\hat{\sigma}}\underline{\mathrm{Tr}(P_i(\textstyle\prod_{s=1}^q \hat{\sigma}_i^{(s)})P_iO_i)}$$

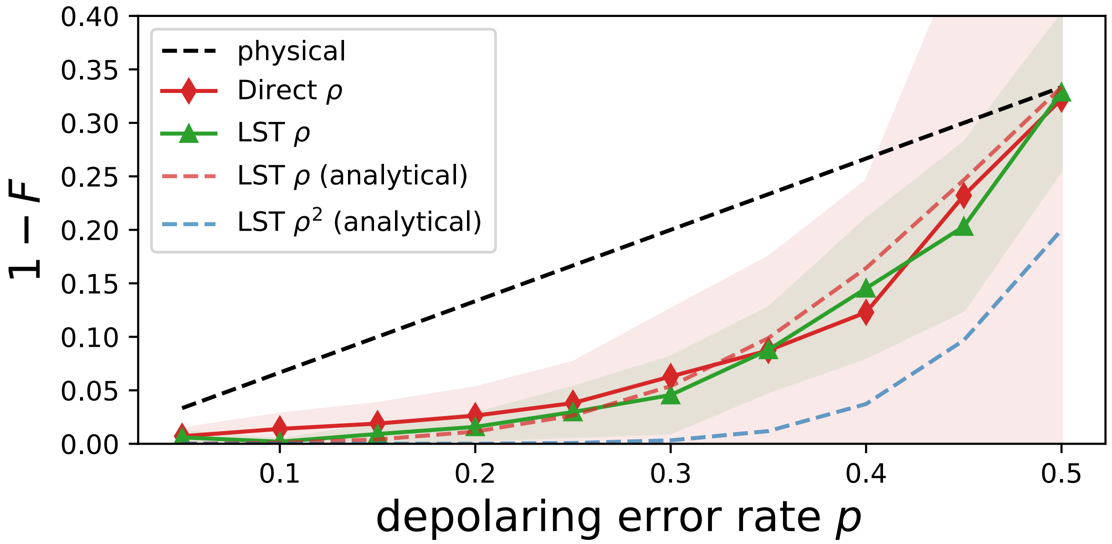
- The evaluation of $\mathrm{Tr}\left(P(\prod_{s=1}^m \hat{\sigma}^{(s)})PO\right)$ is equivalent to finding null space of a binary matrix, which takes $O(n^2m)$ memory and $O(n^3m^2)$ time.

- Logical shadow tomography takes polynomial classical resources in the post-processing.

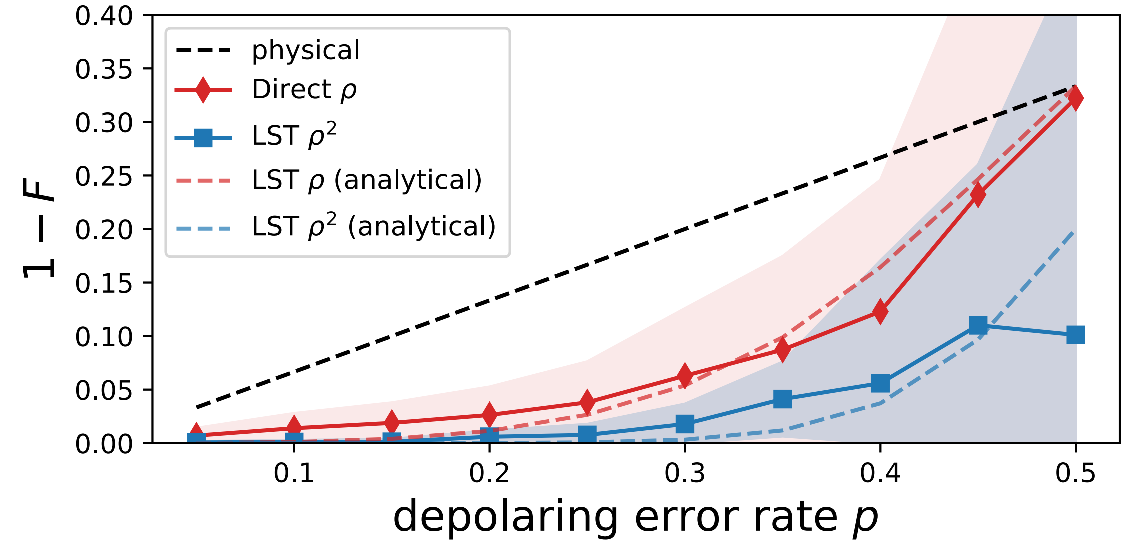
Numerics

- The [[5,1,3]] code:

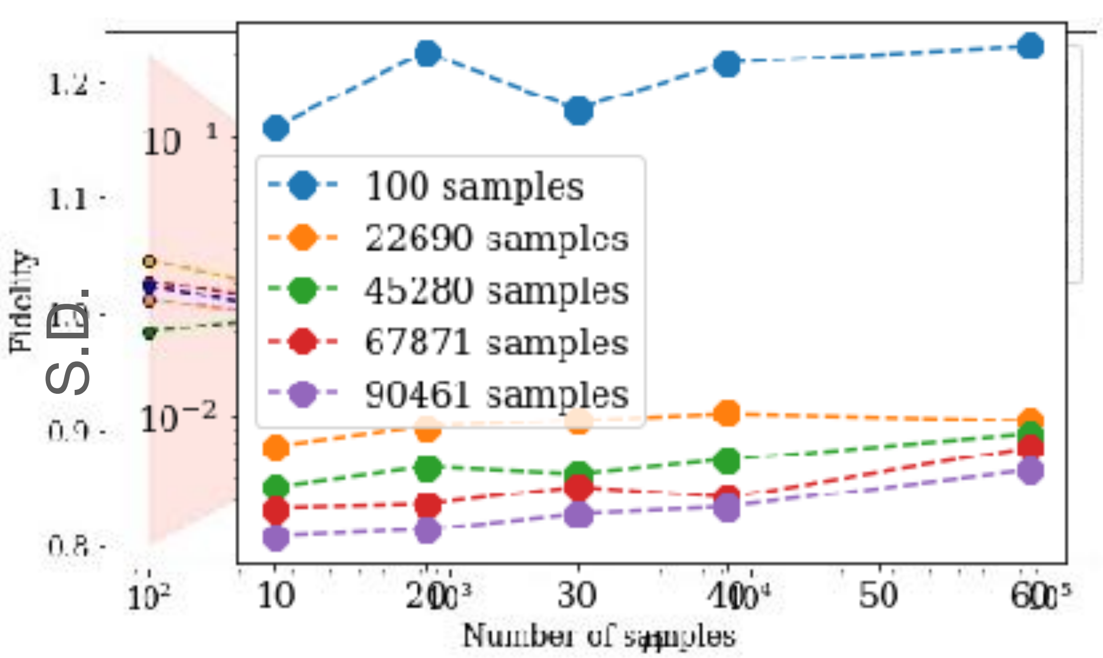
LST with $f(\rho) = \rho$

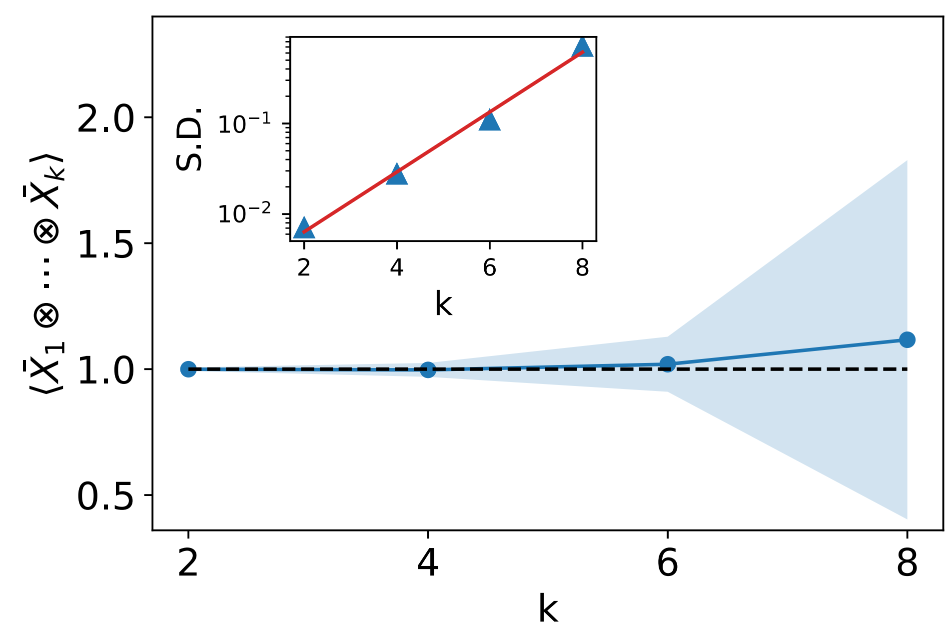


LST with $f(\rho) = \rho^2$



- Compared with SE, LST is more accurate and has better error mitigation
- Sample complexity scaling: $O(4^k)$





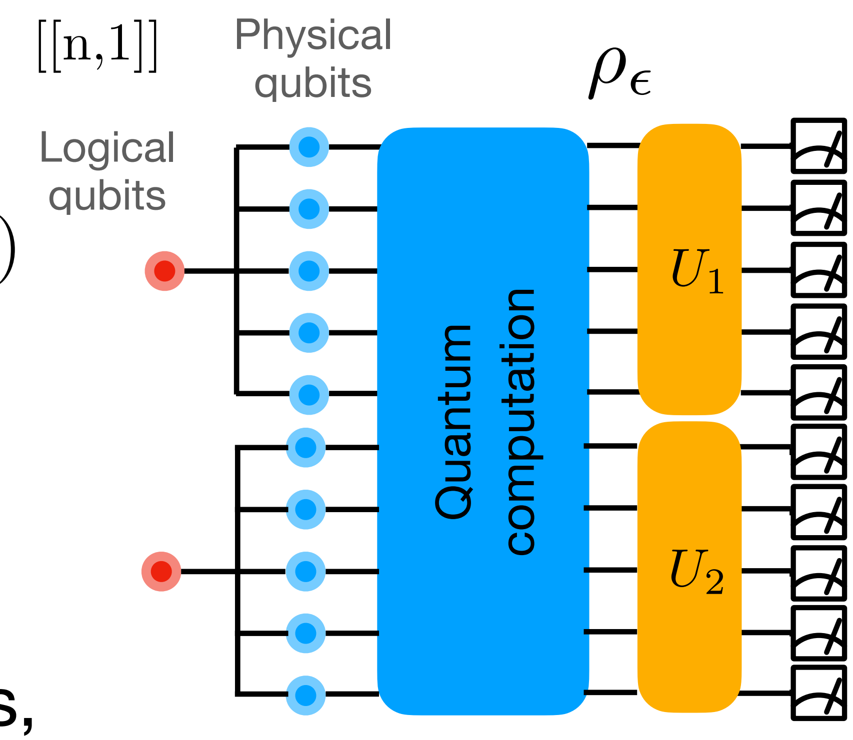
Summary

- Logical shadow tomography:

$$\langle O \rangle_{\mathrm{LST}} = \frac{1}{c}\mathrm{Tr}(Pf(\rho_\epsilon)PO) \quad \text{where} \quad c = \mathrm{Tr}(Pf(\rho_\epsilon)P)$$

  In general, $f(\rho_\epsilon) = c_0 I + c_1\rho_\epsilon + c_2\rho_\epsilon^2 + \ldots$

  SE & VD can be viewed as special cases.



- Low quantum overhead: low depth Clifford circuits, which doesn’t reply on number of logical qubits.

- Better sample complexity: get rid of exponential scaling w.r.t number of physical qubit, and only scales with number of logical qubits

- Polynomial classical resources in the classical post-processing.

Acknowledgement

Hellman Fellows NASA USRA Universities Space Research Association DARPA